Vivekanand Education Society's
Institute of Technology
Approved by AICTE & Affiliated to University of Mumbai

Artificial Intelligence and Data Science Department

Deep Learning/Odd Sem 2023-23/Experiment 2-B

Name : Rupesh Dhirwani	**Class/Roll No. : D16AD - 10**	**Grade :**

Title of Experiment: Implement a backpropagation algorithm to train a DNN with at least 2 hidden layers

Objective of Experiment: The objective of this experiment is to implement a backpropagation algorithm to train a Deep Neural Network (DNN) with a minimum of two hidden layers By doing so, we aim to develop a robust and efficient neural network model capable of leaming and generalizing complex patterns from data, ultimately improving its predictive accuracy.

Outcome of Experiment: We successfully implemented the Practical by successfully training a Deep Neural Network (DNN) with at least two hidden layers using a backpropagation algorithm

Problem Statement: The challenge is to implement a backpropagation algorithm capable of effectively training a Deep Neural Network (DNN) with a minimum of two hidden layers, overcoming issues such as vanishing/exploding gradients, overfitting, and ensuring computational efficiency.

Description / Theory :

Neural Network Basics: Imagine a neural network as a series of connected nodes or neurons, organized into layers: an input layer, one or more hidden layers, and an output layer. Each connection between neurons has a weight, which determines the strength of the connection.

Forward Pass: To make a prediction, we start with some input data. This data is passed through the network in a forward direction. At each neuron, the input is multiplied by the neuron's weights, and then these products are summed up. This sum is then passed through an activation function, which introduces non- linearity into the network. Common activation functions include the sigmoid or ReLU (Rectified Linear Unit) functions. This process of weighted sum and activation is repeated layer by layer, from the input layer through the hidden layers to the output layer, producing the final prediction.

Calculating Error: Once we have made a prediction, we compare it to the actual target value (the correct answer). We calculate the error, which is essentially how far off our prediction is from the actual value.

Backward Pass (Backpropagation): Now comes the magic part. We want to adjust the weights in our network to minimize this error. Backpropagation is the process of figuring out how much each weight contributed to the error and adjusting them accordingly.

Starting from the output layer and moving backward through the layers, we calculate gradients (derivatives) of the error concerning the weights. These gradients tell us the direction and magnitude to adjust each weight to reduce the error.

This is done using the chain rule from calculus, allowing us to distribute the error back through the network layer by layer.

Weight Updates: Once we have the gradients, we update the weights. We move each weight in the opposite direction of its gradient, sealed by a learning rate (a small value). This step ensures that, over time, the network's weights adjust to make better predictions.

Repeat: Steps 2 through 5 are repeated many times with different batches of data (mini-batch gradient descent) to train the network. Each iteration refines the weights, improving the network's ability to make accurate predictions

Training Completion: Training is considered complete when the network's error reaches an acceptable level or after a fixed number of iterations (epochs). In essence, backpropagation is a method for the neural network to learn from its mistakes by iteratively adjusting its internal parameters (weights) until it becomes better at making predictions. This process of forward and backward passes is at the core of training deep learning models.

In essence, backpropagation is a method for the neural network to learn from its mistakes by iteratively adjusting its internal parameters (weights) until it becomes better at maxing predictions. This process of forward and backward passes is at the core of training deep learning models.

Algorithm/ Pseudo Code / Flowchart:

Algorithm:

```
For d in data do
 FORWARDS PASS
        Starting from the input layer, use equation 1 to do a forward pass through the network, computing the activities for the neurons at each layer.
 BACKWARD PASS
        Compute the derivatives of the error function with the respect to the output layer activities.
 For layer in Layers do
        Compute the derivatives of the error function with respect to the inputs of the upper layer neurons.
        Compute the derivatives of the error function with respect to the weights between the outer layer and the layer below.
        Compute the derivatives of the error function with respect to the activities of the layer below.
 End for
 Updates the weights.
End for
```

Program :

```
import numpy as np

X = np.array(([2,9], [1,5], [3,6]), dtype=float)
y = np.array(( [92], [86], [89]), dtype=float)

X = X / np.amax(X, axis=0)
y = y / 100

class NeuralNetwork(object):
    def __init__(self):
        self.inputSize = 2
        self.outputSize = 1
        self.hiddenSize1 = 3
        self.hiddenSize2 = 3
        self.W1 = np.random.randn(self.inputSize, self.hiddenSize1)
        self.W2 = np.random.randn(self.hiddenSize1, self.hiddenSize2)
        self.W3 = np.random.randn(self.hiddenSize2, self.outputSize)

    def feedForward(self, X):
        self.z1 = np.dot(X, self.W1)
        self.a1 = self.sigmoid(self.z1)
        self.z2 = np.dot(self.a1, self.W2)
        self.a2 = self.sigmoid(self.z2)
        self.z3 = np.dot(self.a2, self.W3)
        output = self.sigmoid(self.z3)
        return output

    def sigmoid(self, s, deriv=False):
        if deriv:
            return s * (1 - s)
        return 1 / (1 + np.exp(-s))

    def backward(self, X, y, output):
        self.output_error = y - output
        self.output_delta = self.output_error * self.sigmoid(output, deriv=True)

        self.z2_error = self.output_delta.dot(self.W3.T)
        self.z2_delta = self.z2_error * self.sigmoid(self.a2, deriv=True)

        self.z1_error = self.z2_delta.dot(self.W2.T)
        self.z1_delta = self.z1_error * self.sigmoid(self.a1, deriv=True)

        self.W1 += X.T.dot(self.z1_delta)
        self.W2 += self.a1.T.dot(self.z2_delta)
        self.W3 += self.a2.T.dot(self.output_delta)

    def train(self, X, y):
        output = self.feedForward(X)
        self.backward(X, y, output)

NN = NeuralNetwork()

for i in range(1000):
    if(i % 100 == 0):
        print("LOSS: " + str(np.mean(np.square(y - NN.feedForward(X)))))
    NN.train(X,y)
```

print("Input: " + str(X))

print("Actual Output: " + str(y))

print("LOSS: " + str(np.mean(np.square(y - NN.feedForward(X)))))

print("Predicted Output: " + str(NN.feedForward(X)))
predictions = NN.feedForward(X) * 100

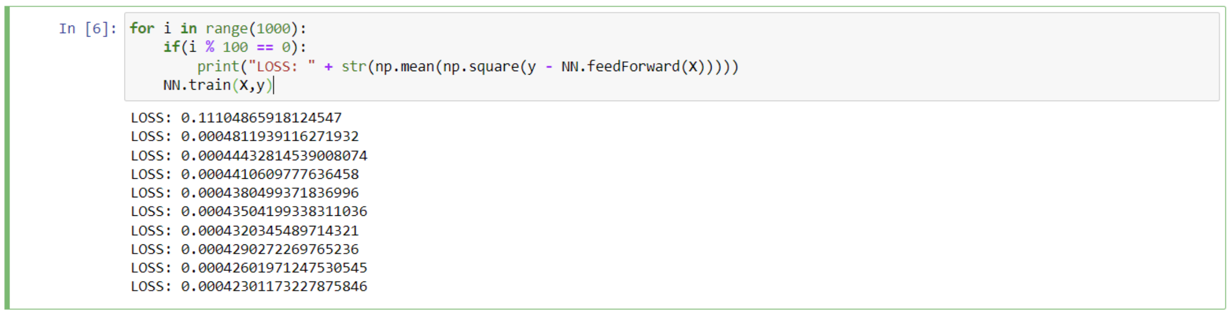
mae = np.mean(np.abs(y - predictions))

rmse = np.sqrt(np.mean(np.square(y - predictions)))

print("Mean Absolute Error (MAE): ", mae)

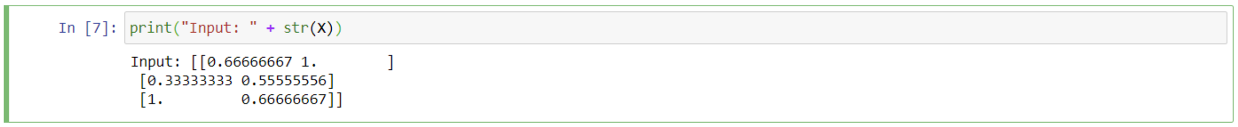
print("Root Mean Squared Error (RMSE): ", rmse)

OUTPUT:



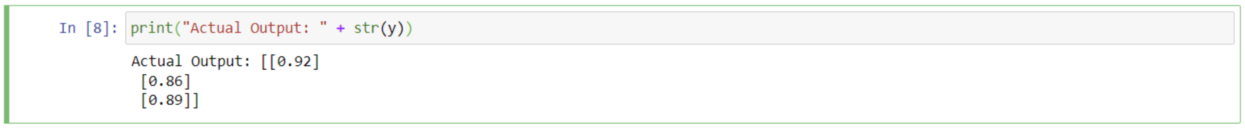
```
In [6]: for i in range(1000):
            if(i % 100 == 0):
                print("LOSS: " + str(np.mean(np.square(y - NN.feedForward(X)))))
            NN.train(X,y)

        LOSS: 0.11104865918124547
        LOSS: 0.0004811939116271932
        LOSS: 0.00044432814539008074
        LOSS: 0.0004410609777636458
        LOSS: 0.0004380499371836996
        LOSS: 0.00043504199338311036
        LOSS: 0.0004320345489714321
        LOSS: 0.0004290272269765236
        LOSS: 0.00042601971247530545
        LOSS: 0.00042301173227875846
```



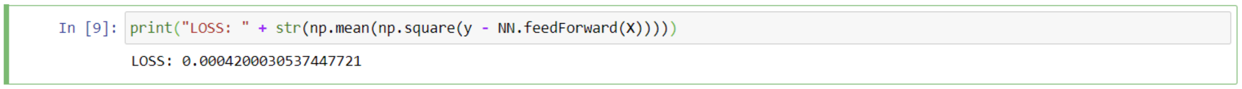
```
In [7]: print("Input: " + str(X))

        Input: [[0.66666667 1.        ]
         [0.33333333 0.55555556]
         [1.         0.66666667]]
```



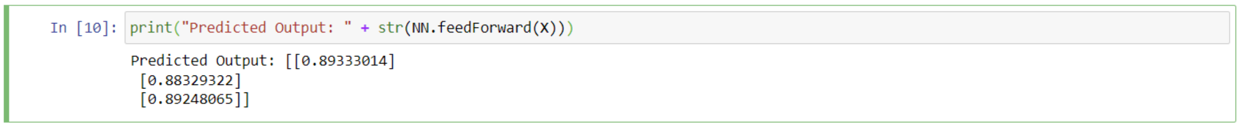
```
In [8]: print("Actual Output: " + str(y))

        Actual Output: [[0.92]
         [0.86]
         [0.89]]
```



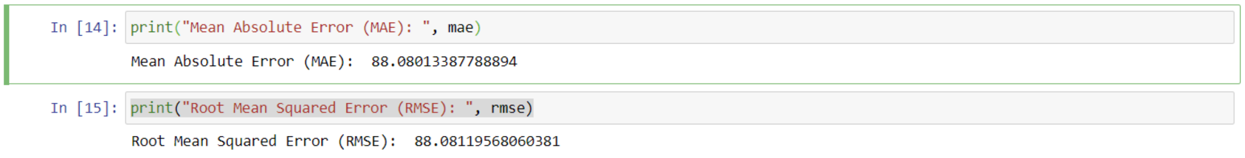
```
In [9]: print("LOSS: " + str(np.mean(np.square(y - NN.feedForward(X)))))

        LOSS: 0.0004200030537447721
```



```
In [10]: print("Predicted Output: " + str(NN.feedForward(X)))

         Predicted Output: [[0.89333014]
          [0.88329322]
          [0.89248065]]
```



```
In [14]: print("Mean Absolute Error (MAE): ", mae)

         Mean Absolute Error (MAE):  88.08013387788894

In [15]: print("Root Mean Squared Error (RMSE): ", rmse)

         Root Mean Squared Error (RMSE):  88.08119568060381
```

Results and Discussions :
We have implemented feed forward neural networks with 3 layers; an input layer with 2 neurons and a hidden layer with 3 neurons each, and an output layer with 1 layer and 1 neuron. The network is trained to predict an output value based on input data.

DATA PREPROCESSING:
Before feeding the data into the network, it is essential to preprocess it. The input features (X) are normalized to values between 0 and 1 using the maximum value of each feature. Additionally, the output values (y) are scaled by dividing them by 100.

Neural Network Architecture: The neural network architecture uses the sigmoid activation functions throughout. It has randomly initialized the weights (W1, W2, W3) for the connections between neurons between each layer.

Training: The training process involves forward and backward passes through the network. In each iteration (1000 iterations in total), the network computes a predicted output and calculates the mean squared error loss. The weights are updated using backpropagation to minimize this loss. The loss decreases with training, indicating that the network is learning to approximate the relationship between the input and the output.

Performance Evaluation: After training the network is tested on the dataset. The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are calculated to evaluate the prediction accuracy. The MAE measures the average absolute difference between predicted and actual values, on the other hand RMSE calculates the measure of overall prediction error. Lower values for these metrics indicate the better performance of the model.

The neural network successfully learns to approximate the values and the relationship between the input data and the target output. The final MAE and RMSE values indicate that the model's predictions are relatively close to the actual values, suggesting that the model has learned the underlying patterns in the data. However, the performance could potentially be improved with the larger dataset, fine-tuning the hyperparameters or using the more complex neural networks architecture. Also the code demonstrates the implementation and training of the basic feed forward neural network for regression tasks.